See alignment included

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12N 15/29, A01H 5/00, 7/00 | A1

(11) International Publication Number: **WO 99/50417**

(43) International Publication Date: 7 October 1999 (07.10.99)

(21) International Application Number: PCT/SE99/00543

(22) International Filing Date: 31 March 1999 (31.03.99)

(30) Priority Data:
9801129-9 31 March 1998 (31.03.98) SE

(71) Applicant *(for all designated States except US)*: A+ SCIENCE INVEST AB [SE/SE]; Kungportsavenyn 31–35, P.O. Box 3096, S–400 10 Göteborg (SE).

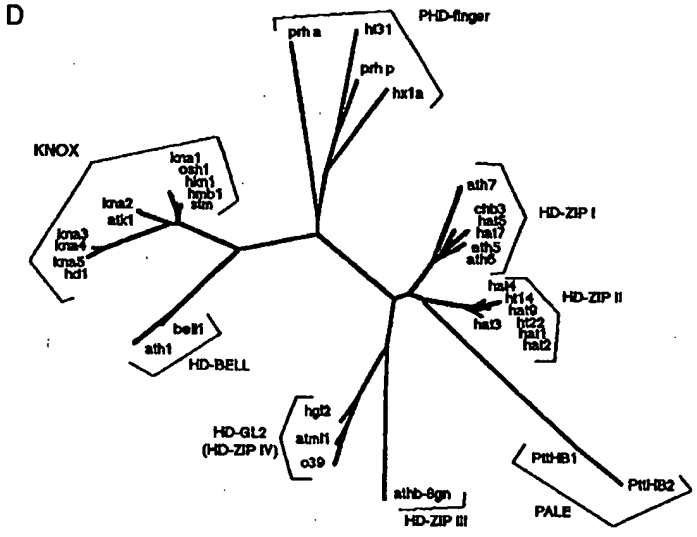
(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: OLSSON, Olof [SE/SE]; Långesandsvägen 33, S–430 90 Öckerö (SE). HERTZBERG, Magnus [SE/SE]; Tvistevägen 46:804, S–907 36 Umeå (SE).

(74) Agent: HOLMBERG, Martin; Göteborgs Patentbyrå Dahls AB, Karlsrovägen 2A, P.O. Box 606, S–182 16 Danderyd (SE).

(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
With international search report.

(54) Title: NOVEL SEQUENCE CLASS OF GENES, CORRESPONDING PROTEINS AND THE USE OF THE SAME



(57) Abstract

The properties of fibrous raw material can be modified already in the growing plant through the functional inclusion of a novel homeobox gene. More specifically, the present invention concerns the expression of homeobox genes in the cambial region of fibrous plants. In particular the present invention concerns the expression of homeobox genes in the xylem maturation zone during secondary phases of vascular development in fibrous plants. A novel class of homeobox genes is disclosed, together with the nucleotide squence and deduced amino acid sequence for five genes from three different plant species, belonging to said sequence class.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

NOVEL SEQUENCE CLASS OF GENES, CORRESPONDING PROTEINS AND THE USE OF THE SAME

The present invention relates to the pre-harvest modification of fibre raw material e.g. the fibrous raw material in the form of fibrous plants. More specifically, the present invention concerns the expression of homeobox genes in the cambial region of fibrous plants. In particular the present invention concerns the expression of homeobox genes in the xylem maturation zone during secondary phases of vascular development in fibrous plants. A novel class of homeobox genes is disclosed, together with the nucleotide sequence and deduced amino acid sequence for five genes from three different plant species, belonging to said sequence class.

Background of the invention

In woody species, the primary growth associated with shoot and root elongation is followed by secondary growth that increases the radial width of xylem and phloem through the activity of the vascular cambium. This meristem is similar to the shoot and root apical meristems in exhibiting highly controlled patterns of cell division, with individual derivative cells having specific developmental fates, but differs from the terminal meristems with respect to origin, position, histology and cytology.

The vascular cambium normally differentiates from the procambium, although it can also arise within a callus. The procambium is a partially differentiated tissue that develops in the embryo and is perpetuated at the shoot apex, where it is initiated in association with the inception of leaf primordia. By convention, the designation of this meristematic zone changes from procambium to vascular cambium in a particular stem portion after it ceases elongating. Thus, within each vascular bundle, there is a gradual basipetal transition from procambium to vascular cambium, and the procambium and vascular cambium are considered to be the same meristem in two stages of development. The procambium-vascular cambium continuum is associated with a gradual change in the characteristics of both the component cells and the derivative xylem and phloem elements. In transverse section, the first derivatives of the procambium are protophloem and protoxylem, which differentiate in the elongating portions of the shoot. Next, metaphloem and metaxylem elements are produced, which differentiate in a shoot portion mainly after it has stopped elongating. Subsequently, derivatives of the vascular cambium are produced, and these cells differentiate into secondary phloem or xylem elements. The vascular cambium that develops within vascular bundles is denoted fascicular cambium. Following its initiation, periclinal divisions occur in the parenchyma cells adjoining

each vascular bundle, apparently induced by a stimulus originating from the fascicular cambium. The resulting interfascicular cambium connects laterally with the fascicular and interfascicular cambia of adjacent bundles, establishing a continuous ring of vascular cambium. The vascular cambium is a highly regulated, dynamic population of partially differentiated cells that can divide in three planes, and whose derivatives differentiate into a variety of genotype-specific cell types comprising two very different tissues, phloem and xylem.

Examination of cell lineages in stem transverse sections of secondary xylem and phloem reveals that individual radial files exhibit changes, such as doubling and disappearance, that occur simultaneously on both sides of the vascular cambium. Such close correspondence can only result from division activity in a common cambial cell, denoted the initial. Following periclinal division in the initial, one of the daughter cells retains the characteristics of the initial while the other one becomes a phloem mother cell or a xylem mother cell, depending on cambium side. The mother cell either differentiates directly or divides periclinally one or more times and all the resulting daughter cells then differentiate.

On the xylem side of the cambium the entire process, from the first division of a cambial initial to the final development of the many possible mature xylem cell types, occurs in several phases. Based on anatomical observations, these phases can be divided into a dividing zone where the xylem mother cells continue to divide, an expansion zone where the derivative cells expand to their final size, a maturation zone where lignification and secondary cell wall thickening occurs, and finally a zone of programmed cell death where all cellular processes are terminated. In trees, little is known about the molecular regulation of this xylogenesis process, which is the bases for wood formation.

Woody plants provide society with materials of major economic importance, e.g., lumber and paper, and, considering the current concern about increasing carbon dioxide levels, represent an important carbon sink. Understanding the regulation of cambial cell division and derivative differentiation will open up possibilities to, by gene technology, alter the developmental fate of the derivatives already during their formation. For example, with this approach it will be possible to modify properties like strength, cell wall thickness, flexibility, homogeneity and surface properties in fibres and vessels of hardwoods and in tracheids of softwoods.

From many developmental studies in animals, insects, worms and also in plants, it is known that several different regulatory circuits interact in complex ways during development. Molecular signals of various kinds are differentially turned on and off, induced

by cell to cell contacts, relative cell positions, environmental cues, nutritional status and/or other long-range stimuli. However, all regulatory steps ultimately work by changing the global pattern of gene expression in an individual cell. This in turn is accomplished in turn by changing the activity of key genes encoding transcription factors, which switch on or off developmental pathways thus triggering a cascade of secondary events and alternate pathways.

There is an urgent need for practical methods of regulating growth speed and the physical and chemical properties of fibrous raw material, as well as for new plants with improved properties in this respect.

Closest prior art

WO 92/17597 discloses recombinant promoters for influencing xylem-specific expression in plants, said promoters preferably derived from the phenylalanine ammonium lyase promoter or homologous to RCR1 or PCR2. However, the disclosed recombinant promoter in itself does not transfer any genetic information regarding the cell differentiation process.

In "Xylogenesis, genetic and environmental regulation – a review" (IAWA Journal, Vol. 17, No. 3, 1996, page 269-310) the author touches the subject of homeobox genes and concludes, that "the regulation of homeobox-gene transcription/translation and consequent production of homeodomain proteins is one mechanism for morphogenetic control." The author further states, that "investigations with putative homeodomain proteins and homeotic genes hold forth hope for resolving the issue of cambial initials."

Apart from such general statements, the present inventors are not aware of any prior art, which would prompt a skilled person in the direction of the present invention, which constitutes a specific and practically feasible approach to the pre-harvest modification of fibrous raw material. The problem underlying the invention and presented under the heading Background of the invention, remains unsolved.

Summary of the invention

The above stated problem is solved through the invention as disclosed in the attached claims. The present investigators have identified key genes or *de facto* an entirely novel class of such genes and put these into practical use. The present invention discloses a novel class of homeobox genes, distinguished from previous known classes, influencing the

cell differentiation and growth of fibrous plants. The present invention further makes available novel transgenic plants and technical methods for their production.

Short description of the drawings

The present invention is described in further detail below with reference to the enclosed examples and figures, in which:

Fig. 1 shows the DNA sequence and deduced amino acid sequence of PttHB1. Amino acids are given in the one-letter code under the nucleotide sequence. A number indicates the start and stop of translation. Residues representing the homeodomain (HD) are boxed. Possible open reading frames on the 5' leader sequence preceding the initiation codon are indicated in italics, and putative nuclear localisation signals are underlined.

Fig. 2 shows amino acid sequence alignments of PttHB1 and PttHB2 HD with each other and with HD sequences from diverse organisms. Horizontal bars above of the sequences indicate the positions of the three helixes in the HD. Identical amino acids are shaded black, similar residues are shaded grey, and non-conserved amino acids are not shaded. Gaps in the sequence are indicated by dots.

A. PttHB1 and PttHB2 HD. Very conserved aa are indicated by filled triangles, and core hydrophobic aa by an H (see text for explanation)
B. Closest HD from any organism
C. 34 chosen plant HD, representing all 4 plant HB classes
D. Phylogenetic tree of HD from the same 34 plant HB genes, plus the hybrid aspen HB genes
E. Alignment of the entire PttHB1 aa sequence with corresponding sequence from PttHB2, AtPALE1, ATPALE2 and PrPALE2 HD.

(Accession numbers: DANF, sptrembl|P79738|; PRX2_chicken, sptrembl|O90963|; hat9_arath, swiss P46603; ht22_arath, swissP46604; ht14_arath, swissP46665; hat1_arath, swiss P46600; hat2_arath, swiss P46601; hat4_arath, swiss Q05466; hat3_arath, swiss P46602; chb3_carrot, pironly|s51927|; hat5_arath, swissQ02283; ath5_arath, swiss P46667; ath6_arath, hat7_arath, swiss Q00466; ath7_arath, swiss P46897; atml1, EMBL AT37589; o39, EMBL U34743; hgl2_arath, swiss P46607; athb-8GN, EMBL Z50851; bell1, EMBL U39944; ath1_arath, swiss P48731; kna3_arath, swiss P48000; kna4_arath, swiss P48001;

kna5_arath, swiss P48002; hd1_brana, swiss P46606, atk1, EMBL X81353; kna2_arath, swiss P46640; stm, EMBL AT32344; hmb1_soybn, swiss P46608; hkn1_maize, swiss P24345; osh1_orysa, swiss P46609; kna1_arath, swiss P46639; ht31_arath, swiss Q04996; prh_petcr, swiss P48786; hx1a_maize, swiss P46605; prh_arath, swiss P48785.)

Fig. 3 shows a Southern hybridisation experiment to demonstrate the presence of *PttHB* like sequences in the plant kingdom.

(A) Chromosomal DNA isolated from hybrid aspen was digested with restriction enzymes as indicated, and the same filter was probed with either a 3' region *PttHB1* probe or a full length *PttHB2* probe. The filters were hybridised and washed at stringent conditions.

(B) To the left the same filter as in Figure 4A. To the right, a filter containing chromosomal DNA from maize, *Arabidopsis*, tobacco, coffee and Norway spruce, digested with *Eco*RI. Both filters were hybridised and washed at a less stringent condition compared with (A) using a full-length *PttHB1* probe.

Abbreviations: E; *Eco*R1, B; *Bam*H1, H; *Hind*III and Ev; *Eco*RV.
Zm; *Zea maize*, At; *Arabidopsis thaliana*, Nt; *Nicotiana tabaccum*, Ca; *Coffea arabicum*, Pa; *Picea abies*.

Fig. 4 shows a northern hybridisation experiment to demonstrate *PttHB1* and *PttHB2* expression in hybrid aspen plants. Total hybrid aspen RNA, isolated from different plant tissues as indicated, was hybridised to either a full length *PttHB1* probe, a full length *PttHB2* probe or an actin probe from *P. trichocarpa* under stringent conditions. Estimated sizes in bases of the detected transcripts, calculated from size markers co-run with the RNA samples, are indicated to the right.

R; Root, X; Differentiating xylem, P; Differentiating phloem, B; Bark and L; Leaf.

Fig. 5 shows *PttHB1* and *PttHB2* expression in the hybrid aspen stem.
(A) Nomarski optics microscope picture showing the different developmental zones in the cambial region of a hybrid aspen stem. Tissue samples isolated by sectioning and used for mRNA isolation are indicated at the bottom as C1; C2; C3 and C4. These sample zones correspond to differentiating phloem (Ph), cambial zone (Cz), enlarging xylem (EZ) and maturing xylem (MZ), respectively. Horizontal bars indicate the length of the developmental and sectioned zones, respectively.

(B) *PttHB1* and *PttHB2* expression in the cambial region of hybrid aspen, as reflected by PCR amplification of mRNA, isolated as depicted in Figure 5A, and analysed by Southern hybridisation. The *PttHB1* and *PttHB2* probes used were the same as in Figure 4.

Description of invention

Numerous studies have demonstrated crucial roles of homeobox (HB) genes in the control of a vast diversity of cellular and developmental processes, such as spatial patterning, positional information, cell fate determination and cell differentiation, in eukaryotic organisms from yeast to man. Since the first discovery of HB genes, the number of genes identified which carry a HD motif has steadily increased, emphasising the great importance of this gene class in all biological systems studied.

The HB itself is a semi-conserved DNA sequence of about 180 base pairs (bp) found in the coding region of HB genes, encoding a 60 amino acid (aa) homeodomain (HD) motif. NMR and X-ray crystal structures of several HD domains have been determined. From these studies it can be concluded that although the primary HD aa sequence can be quite divergent among different genes, the secondary structures are remarkably similar, consisting of a flexible N-terminal arm followed by a helix-loop-helix-turn-helix structure. Therefore, different HD domains most likely have a very similar three-dimensional structure.

If a 55-60% similarity is used as a criterion, HD sequences can be grouped into at least 30 distinct classes. Some of these classes have been placed into common superclasses, as HEX, PRX and TALE. Furthermore, some HB classes are further divided into families. The most important criterion for designating a novel gene to a HB gene superclass, class or family is the structure of the HD itself, due to its important functional implications, mainly in the protein/DNA interaction. In many instances, however, domains outside the HD are conserved and are also used in the designation of the individual HB genes. Also in plants, the list of described HB genes is rapidly increasing. These genes are rather diverged, and presently fall into four different HB classes namely HD-ZIP, HD-KN (KNOX), HD-BELL, and PHD-finger (Figure 2C). The HD-ZIP class has been further subdivided into four families, named HD-ZIP I, II, III and IV.

The two HB genes identified by the present inventors were denoted *PttHB1* and *PttHB2*. When aligned to members of all known HB gene classes, the *PttHB* HDs only showed a 47% aa identity to the closest HD, the *DANF* gene from Zebra fish. In addition the deduced proteins encoded by the *PttHB* genes do not contain any other characteristic motifs outside the HD that are reminiscent of any other HB gene, and they show a unique 5 aa loop

extension between helix one and two. The 5 aa insertion puts the *PttHB* sequence into the particular subset of HB genes that have more or fewer than 60 aa in the HD. Several previous studies have shown that such aa are accommodated either between helix 1 and 2, or helix 2 and 3. A new superclass of HB genes has recently emerged from this subset, denoted TALE for three amino acid loop extension. As the name indicates, gene members of this superclass, has three extra aa inserted between helix 1 and 2 in the HD. The plant HB classes *KNOX* and *BELL* for example, fit into this superclass. However, an extension of five aa has not previously been seen in any HB gene. The present inventors therefore suggest that the *PttHB* genes are the first members of a new class, which is tentatively named PALE, for penta aa loop extension. It is possible, as more genes belonging to this class are found that also PALE will emerge as a superclass.

When the PttHB aa sequences were aligned to the HD domain of 34 HB genes from plants, including members of all four HB classes it became clear that the closest plant gene, *HAT2* belonging to the HD-ZIP-II class, have an even lower identity to the PttHB aa sequences than PRX2, or 35% (Figure 2B and 2C). In addition, when the HD in the *PttHB* genes were compared to the same plant genes in an evolutionary analysis it became obvious that the hybrid aspen *PttHB* genes form a HB plant class of their own (Figure 2D). This class is thus evolutionarily distinct from the four previously described plant HB classes, which also can be distinguished in Figure 2D. The relative evolutionary distances between the *PttHB* sequences and previously described plant HB classes was confirmed also when protein sequences outside the HD (Figure 2E) were included in the analyses (data not shown). Accordingly, the present inventors conclude that the hybrid aspen *PttHB* genes are the first described examples of a new class of HB genes. This class has a 5 aa loop extension typically containing the aa QKIK motif, between helix 1 and helix 2, an ITXE motif in helix 2, breaking this structure somewhat and sometimes a WTP motif in the N-terminal arm of the HD.

Searches for HB genes with 5 aa loop extensions between helix 1 and helix 2 in the genomic databases, and PCR experiments using primers from the PttHB2 sequence revealed that indeed members of the PALE class, fulfilling the above criteria, are present in *Arabidopsis* and *Pinus radiata* (Fig. 2E).

Expression studies revealed that the two hybrid aspen *PttHB* genes were differentially regulated (Figures 4 and 5). The *PttHB1* gene displayed a tissue-specific expression, being active in the xylem maturation zone of the cambial region. The *PttHB2*

gene on the other hand, was active in earlier developmental phases on both sides of the cambium, as well as in the cambium itself (Fig. 4 and 5).

The present inventors have surprisingly identified, isolated and characterised HB like cDNA sequences isolated from a cambial cDNA library from the hardwood *Populus tremula* x *tremuloides*. In addition they have isolated similar HB like DNA sequences from the softwood *Pinus radiata* and the annual model plant *Arabidopsis thaliana*. These cDNAs and sequences do not fall into any previously described HB sequence class from any system, thus making up a novel HB sequence class of their own. Furthermore, this sequence class is evolutionary more distant to plant than to animal HB genes. One of the cDNAs is specifically expressed at the stage of xylogenesis where secondary fibre wall formation is initiated, and present data indicate that this cDNA is involved in the regulation of this secondary development. The implications of these findings in relation to molecular regulation of wood formation are far-reaching.

Examples

Materials and methods

Plant material

Material was harvested from hybrid aspen (*Populus tremula* x *tremuloides*) plants that were 1.5-3 m tall. They were grown in the greenhouse under natural light supplemented with metal halogen lamps, giving a photon flux density of 150 µmol/m^2/sec, a photoperiod of 18 h, and at a temperature of about 22/15°C (day/night). The plants were watered daily and fertilised once a week with a complete nutrient solution containing 100 mg nitrogen per litre. Material used for library constructions, however, was isolated from plants grown under more standardised conditions, using a controlled environment chamber with a photon flux density of 240 µmol/m2/sec (Osram HQI-TS 400 W/DH metal halogen lamps), a photoperiod of 18 h, a temperature of 20/10°C (day/night), and a relative humidity of about 70%. These plants were also watered with a complete nutrient solution containing 100 mg nitrogen per litre.

Plant tissue culture and genetic transformation

To obtain the starting sterile tissue culture material for genetic transformation, root segments of the hybrid *Populus tremula* x *P. tremuloides* were buried in moist peat, and sprouts were induced under greenhouse conditions (natural photoperiod extended to 18 h as required by metal halogen lamps giving a photon flux density of 150 µmol/m^2/sec, a day/night temperature of 23/16°C, and a relative humidity of at least 50%. Shoots 100 to 200-

mm-tall were surface-sterilised for 10 min in 0.1% $HgCl_2$ and rinsed three times in sterilised water. Segments 15-mm-long were excised from the stem, avoiding the nodes where possible, and placed on solid medium, hereafter referred to as MS1, which contained 0.1 µg/ml indole-3-butyric acid (IBA), 0.2 µg/ml 6-benzylaminopurine (BAP), and 0.001 µg/ml thidiazuron (TDZ; N-phenyl-N-1,2,3-thidiazol urea) to initiate shoots. The cultures were grown in a controlled environment room having a temperature of 25°C, a photoperiod of 16 h, and a light intensity of 40 $\mu E/m^2/sec$ from cool white fluorescent lamps. When the shoots were about 5 mm long, the cultures were transferred to MS2 medium (MS1 medium minus TDZ) to promote shoot elongation. After their length exceeded 6 cm, the shoots either were used for transformation (see below) or were induced to root by placing them on MS3 medium (1/2 strength MS medium without hormones). Rooted shoots were potted in peat, covered with a plastic bag, and placed in the greenhouse. The bag was ventilated one week later, and removed after a second week.

The DNA was introduced into the plants by *Agrobacterium* mediated transformation. Fresh cultures of *A. tumefaciens* cells were made electrocompetent by growth in (yeast-extract beef) YEB medium (0,1% yeast extract, 0,5% beef extract, 0,1% peptone, 0,5% sucrose and 2 mM $MgSO_4$) to an OD_{595} of 0.5, washed three times in distilled water, resuspended to about 10^9 cells/ml in 10% glycerol. Competent cells were stored at –70°C.

Cells were thawed on ice, and a 50µl aliquot was mixed with 50 ng of vector DNA. A single pulse (Gene Pulser, BioRad) was delivered to the mixture at 2 kV, 25 µF and 200 ohms. The electroporated cells were immediately transferred to recovery medium (YEB supplemented with 10 mM NaCl, 2,5 mM KCl, 10 mM $MgCl_2$ and 10 mM $MgSO_4$). After incubation at 28°C for 2 h, transformed cells were plated on solid YEB medium containing 100 µg/ml rifampicin, 100 µg/ml carbenicillin and 25 µg/ml kanamycin, and incubated at 28 °C for 24 to 48 h. Single colonies were restreaked on fresh selective medium.

To check the integrity of the transferred binary vector, a back-cross to *E. coli* was made. Single *Agrobacterium* colonies were grown to an OD_{595} of 0.5. An over-night culture of *E. coli* strain DH5a was spread on LA plates containing 100 µg/ml carbenicillin. The plates were allowed to dry for 5 min, after which the fresh *Agrobacterium* cells were spotted onto the lawn of *E. coli* cells. The plates were incubated at 28°C for 6 h, then moved to 37°C over-night. Small colonies of *E. coli* cells then appeared inside the *Agrobacterium* spots. These *E. coli* colonies were re-streaked on LA plates containing carbenicillin. Plasmids were isolated from these cells and physically mapped by restriction enzyme analysis.

Agrobacterium cells carrying binary plasmids containing PALE HB genes in

sense and anti-sense directions were grown in YEB medium supplemented with 50 µg/ml carbenicillin and 25 µg/ml kanamycin at 25°C for about 24 h. When OD_{595} reached 0.2-0.6, the cultures were centrifuged for 10 min at 3000 rpm, resuspended in MS medium containing 20 µM acetosyringone and grown at 28°C for one hour on a gentle reciprocating shaker. Acetosyringone was applied to increase the efficacy of gene transfer between the *Agrobacterium* and the plant cell.

Stem segments of hybrid aspen were co-cultivated with *A. tumefaciens* cells in liquid MS medium for 0.5-2 hours, then transferred to MS1 medium. After incubation for 48 h in the dark, the segments were washed twice in sterile 500 µg/ml cefotaxime and placed in the light on MS1 medium supplemented with 250 µg/ml cefotaxime and either 60 µg/ml kanamycin or 15 µg/ml hygromycin, depending on the vector used for gene transfer. After shoot initiation, the segments were transferred to MS2 medium, containing the same antibiotics, to promote elongation. Roots were initiated by transferring the cultures to MS3 medium. Rooted shoots were transferred to the greenhouse, and potted in 1.5 l pots filled with fertilised peat. After the peat was thoroughly watered the plant was covered with a plastic bab. During acclimatisation, direct exposure to artificial light or sunlight was avoided. After about 1 week, the plastic bag was ventilated and then removed after about 2 weeks. The timing of the acclimatisation procedure was adjusted to the vigour of the plantlet in every instance.

Preparation and screening of a cDNA library

Cambial region material, containing cells from the phloem, the cambial zone and differentiating xylem, was collected by peeling the bark and scraping the inside of the bark peeling and the outside of the exposed xylem with a scalpel. Both scrapings contained fibres. Poly (A) RNA was isolated from cambial scrapings by means of magnetic oligo (dT) beads (Dynabeads® Oligo (dT) 25 Dynal A.S., Oslo, Norway), according to the manufacturer's recommendations. A λgt22a cDNA library was constructed (Superscript™ Lambda System for cDNA Synthesis and Cloning, Gibco BRL, Gaithersburg, USA) and packed into λ-particles, again according to the manufactures instructions (Gigapack II Gold, Stratagen, La Jolla, USA). *E. coli* Y1090r was used as a bacterial host. The complexity of the library obtained was 900 000 pfu. The library was amplified once on plates and 200 000 pfu of the amplified library were screened by plaque hybridisation with a degenerate oligonucleotide, denoted HB2, (5'-TGG TTY CAR AAY MGN MG-3') which recognises the conserved helix 3 of homeobox genes. The oligonucleotide was 5' end-labelled with T4 polynucleotide kinase using [γ-^{32}P] ATP, 5000Ci/mmol (Amersham) (Gibco BRL) and purified on a Sephadex-G50

column (Pharmacia, Sollentuna, Sweden). Plaque blotting was performed as described by the manufacturer (Hybond-N, Amersham). Filters from the first screen were washed for 5 x 20min at room temperature in 6 x SSC + 0.05% sodium pyrophosphate, while filters from the 2nd and 3rd screens were washed in the same buffer as described by Bürglin et al, but at 50°C. DNA from purified phages was PCR amplified by means of universal primers (Lambda gt11 Forward and Reverse, Promega, Madison, USA) flanking the cDNA insert. PCR products were purified on gel and sequenced with the Lambda gt11 Forward primer.

DNA sequencing and subcloning

The two cDNAs chosen were subcloned into the *Not*I/*Eco*RI and *Not*I/*Sal*I sites, respectively, of the cloning vector pOK12. All cDNAs presented were sequenced on both strands by the dideoxynucleotide chain termination method using the ABI PRISM system (Perkin Elmer, Warrington, Great Britain).

Southern hybridisation

Chromosomal DNA from all plant species investigated was isolated from young leaves. Ten µg of genomic DNA were digested with *Eco*RI, *Bam*HI, *Hind*III or *Eco*RV, separated on an 1% agarose gel, and blotted to nylon filters (Hybond-N) using a vacuum blotting device (VacuGene XL, Pharmacia LKB, Sweden). In a second Southern blot experiment, 15µg of DNA from Norway spruce, 5µg from coffee, 5µg from corn and 2µg from *Arabidopsis thaliana* were digested with *Eco*RI and treated as above. Probes were isolated as a 620bp *Sac*I/*Not*I *PttHB1* 3' fragment, a 1100 bp *Not*1/*Eco*RI full length *PttHB1* fragment and 1200 bp full length *Not*I/*Sal*I *PttHB2* fragment, respectively. They were labelled by γ^{32}P-dATP, using the random labelling reaction. Southern hybridisations were performed in Church buffer at 65°C, or alternatively at 50°C for low stringency. Final washings were performed in 0.1xSSC at 65°C or in 2xSSC at 50°C for low stringency. The radioactivity on the filters was finally analysed on a phosphor-imaging system (GS-525 Molecular Imager®, Storage Phosphor Imaging Systems, BioRad, Solna, Sweden).

RNA isolation and northern hybridisation

Samples for RNA isolation were collected from young plants about 3 m tall by peeling the bark and scraping the inside of the bark peeling to obtain cambial zone cells + differentiating phloem (denoted the phloem fraction) and the outside of the exposed xylem to obtain differentiating xylem (denoted the xylem fraction). The tissue that was left after

scraping the phloem was considered as bark. In addition, a leaf sample was taken from young, but fully expanded leaves after the mid-vein had been removed. Finally, a sample of young roots was also collected. Twenty-five µg total RNA from each sample was separated on an glyoxal gel and blotted to a nylon filter (Hybond-N). All subsequent hybridisations were performed as described above in "Southern hybridisation".

High resolution expression study

Samples from the different zones of xylem development were obtained from a hybrid aspen stem segment by longitudinal tangential sections through the cambial region using a cryo-microtome (HM 505 E microtome, Microme Laborgeräte, Walldorf, Germany). Transverse hand sections were taken at the same time to elucidate the location of each section. One of these transverse sections is shown in Figure 5A. The different developmental zones were assigned as Ph (Phloem), CZ (Cambial zone), EZ (Enlarging xylem zone) and MZ (Maturating xylem zone). The distinction of the different zones was made based on the radial diameter of the cells and the presence of birefringence, as seen under the light microscope using Nomarski optics. Individual 50 µm thick, 2.5 x 15mm sections were pooled into developmental groups as indicated by C1 (differentiating phloem), C2, C3 and C4 in Figure 5A. Poly (A) RNAs were isolated from each tissue using magnetic oligo (dT) beads. First-strand cDNA was synthesised on the beads using a first-strand cDNA synthesis kit (First-Strand cDNA Synthesis Kit, Pharmacia Biotech), free poly (dT) was removed from the beads by T4 DNA polymerase (BRL), and RNaseH (MBI Fermenta Sweden) was used to remove RNA in the cDNA/RNA hybrid. A homopolymeric (dA) tail was added to the ss cDNAs using terminal deoxy transferase TdT. Finally, the synthesised cDNAs were amplified by PCR. This was done by adding 50µl PCR reaction mixtures to the cDNAs (200µM dNTPs, 1.5 mM MgCl2, 10 mM Tris buffer pH 9.0, 1U Taq DNA polymerase and 0.5µM *Xba*I-(dT)17 primer (5'-GCGCCATCTAGAGCTTTTTTTTTTTTTTT-3'). Initially, a cycle of 1 min at 94°C, 2 min at 40°C and 3 min at 72°C was run 10 times. After this, the beads were removed from the reaction mixture, and 40 more cycles (1 min at 94°C, 2 min at 58°C, and 3 min at 72°C) were run. The amplified cDNAs were separated on an agarose gel and the relative amount of each PCR product was estimated after staining with ethidium bromide. Similar amounts of each amplification product were loaded on 1% agarose gel, separated and blotted to a nylon filter, and hybridised to the *PttHB1* and *PttHB2* probes as described above in "Southern hybridisation".

Microscopy

Microscopy on transverse hand sections was done on an Axioplan (Carl Zeiss) microscope using Nomarski optics. The specimens were mounted in 100% glycerol.

DNA sequence analysis

The Gene Construction Kit (Textco Inc., West Lebanon, New Hampshire, USA), and MacVector 4.5 (Scientific Imaging System, New Haven, CT, USA) soft ware was used for visualising constructions and sequences, for analysing sequence data and for local aligning of various DNA sequences. DNA sequence similarity searches were performed in Basic Local Alignment Search Tool 2 (BLAST2) directly on line to EMBL, Heidelberg, Germany. The translation products of the homeobox genes described here were aligned to other homeobox translation products using the program Pileup in the GCG package (Genetics Computer Group, Wisconsin, USA). Phylogenetic analysis to create the tree were done on hand-modified PILEUP alignments using the GCG programs DISTANCES and GROWTREE, the Kimura Protein distance matrix and the neighbour-joining method. The phylogenetic tree was plotted with TREEVIEW PPC.

Fibre cell measurements

Tracheids and vessels were isolated from wt and transgenic hybrid aspen plants by maceration in H_2O_2 and acetic acid at 100°C for 4 h. Fibre cell length and cell wall thickness (CWT) were measured in a Kajaani FiberLab apparatus and recorded as relative Kajaani units.

Results

Nucleotide sequences of the PttHB1 and PttHB2 cDNAs

To elucidate the role of HB genes in plant vascular development, the present inventors synthesised a primer mix corresponding to very conserved 8 aa residues from the third helix region of the homeodomain (HD). This mix was subsequently used as a probe in hybridisation experiments. The present inventors screened 200 000 plaques of a cambial region library, and found 26 cDNA clones, indicating a relatively low expression of HB genes in hybrid aspen vascular tissue. DNA sequence information from 12 of these clones showed that they originated from two different genes, which were designated *PttHB1* and *PttHB2* (for *Populus tremula* x *tremuloides* homeobox). The entire nucleotide sequence of the *PttHB1* cDNAs is presented in Figure 1. This cDNA sequence is 1094 bp long, excluding the poly (A)

sequence. It contains an open reading frame (ORF) of 651 nucleotides, starting with an ATG codon at position 136 and ending with a TGA codon at position 787. The *PttHB1* coding sequence is preceded by a 135 bp AT-rich 5' untranslated region, followed by a 304 bp untranslated 3' sequence and ending with a poly (A) tail.

The immediate context around the presumed translational start codon is GCTCATGGA, which is sub-optimal if compared to a derived translational initiation consensus sequence, caAA/CaATGGCg, for plants. However, the most important position is believed to be the G at position +4 (where A in ATG is defined as +1), which is present in the *PttHB1* initiation site. The relatively poor AUG context in the *PttHB1* transcript could reflect that this gene encodes a transcription factor and thus a less efficient translation in fact results in optimal cellular levels of this particular protein. In all other respects however, the *PttHB1* transcript looks like a typical plant gene transcript. It has a relatively short and AU rich leader sequence, which reduce the potential for secondary structure formation and no additional in-frame ATG codons can be found in the sequence upstream of the presumed starting ATG. One 2-aa-long ORF is present in the 135 bp sequence preceding the *PttHB1* coding sequence. The significance, if any, of this ORF is not known at present, but it could possibly be involved in translational regulation of the *PttHB1* gene. No clear consensus sequences for polyadenylation signals (AATAAA/T) is present immediately upstream of the poly (A) insertion site, although a poly (A) tail is clearly present in the cDNA. However, in many plant genes such signals appear diffuse, and it has not been possible to define a single, universal poly (A) signal. Putative nuclear localisation signals are present at aa positions 130 - 137 and 158 – 163, indicating that the PttHB proteins mediate their biological activity in the nucleus.

Predicted amino acid sequence of the PttHB1 gene

The predicted PttHB1 protein is 217 aa long, which gives a calculated molecular weight of about 24 kDa. Between positions 352 and 550 in the DNA sequence a peptide motif corresponding to the consensus HD is found. In the *Antp* HD the tertiary structure is held together by a hydrophobic core of twelve tightly packed aa. All but one of these aa are hydrophobic, and six are highly conserved or invariant. In the PttHB1 HD, eight out of the eleven conserved hydrophobic aa are present. At the other three positions two hydrophobic aa have been replaced by two others, and the single non-hydrophobic aa in the core is basic. An H in Figure 2A marks these eleven structurally important aa. The missing core aa should be situated between helix 1 and 2, but since the PttHB1 aa sequence has a five residues insertion

in this region, it is not possible to unambiguously define the position of this last hydrophobic aa.

The PttHB1 HD contains all seven of the most highly conserved aa at positions where only one or two aa has been found in all the different species analysed so far, and which serves as a signature for HD domains. These aa are marked in Figure 2A as black triangles.

In the N-terminal arm of HD domains, two arginine residues at positions 3 and 5 directly contact the DNA in the minor groove. These two residues are also present in the PttHB1 HD (Figure 1, 2A). Further contacts between HD proteins and target DNA are by the recognition helix III aa residues I.QNRRM/A, touching the major groove of the DNA around a TAAT motif. In the PttHB1 HD domain aa residues QNRRA are indeed present (Figure 2A). Thus. despite the sequence divergence between the PttHB1 HD and other known HD classes (Figure 2B, 2C), the general three-dimensional structure of the HD seems to be conserved. No other conserved sequence motifs, found on other homeobox genes outside the HD, are present.

The *PttHB1* HB sequence is more similar to the *HAT2* sequence than to other plants HB. The *HAT2* gene belongs to the HD-ZIP II family. However, no leucine zipper motif is present in the *PttHB1* sequence excluding this gene from the HD-Zip plant class of HB genes. In addition, no other signatures of interest, such as acidic regions with clusters of aspartic and/or glutamic acid, common in many transcription factors, can be found. However, some homology to the herplex simplex virus strong trans-activating domain of VP16, which functions in a large number of biological systems, including plants is found between aa 152-203.

The PttHB genes represent a novel homeobox class

The two PttHB aa sequences were aligned to each other, as well as to HD regions of representatives of various classes originating from many species of animals and plants. This showed that the *PttHB1* and *PttHB2* genes encode a HD with a length of 65 aa, due to 5 extra aa inserted between helix 1 and 2. The present inventors used these 5 aa as a distinguising mark for this novel HD class, denoted PALE for Pentra AA Loop Extension. As depicted in Figure 2A and E, the PttHB aa sequences, as well as the *A. thaliana* AtPALE 1 and 2 aa sequences and the *Pinus radiata* P.rPALE 1 aa sequence are very similar to each other. showing an about 75% identity over the HD. and an overall identity of about 60% over

the whole translated proteins. On the other hand, none of these PALE sequences were found to be very close to any other HB sequences.

The PttHB genes are part of a small gene family.

Southern blot analysis of hybrid aspen genomic DNA at high stringency showed that the shorter 620 bp *PttHB1* 3' probe hybridised to a single band in the *Hind*III and *Eco*RV genomic digests and to two or three bands in the *Eco*RI and *Bam*HI digest (Figure 3A). When using a full length *PttHB2* probe, a single band was obtained in the *Eco*RV digests, and two to three bands in the *Eco*RI, *Bam*HI and *Hind*III digest (Figure 3A). This indicates that both *PttHB1* and *PttHB2* most likely are single-copy genes, and that the multiple bands obtained in some of the digests are due to introns interrupting the sequence corresponding to the respective probe. When using a full-length *PttHB1* probe in a low stringency hybridisation to genomic DNA from maize, *Arabidopsis*, tobacco, coffee and spruce, multiple bands were obtained (Figure 3B). This showed that *PttHB1*-like sequences are present in maize, tobacco and Norway spruce, indicating that this new HB class is widespread in nature. At the stringency used however, no distinct bands were obtained from *Arabidopsis* or coffee DNA.

The PttHB1 and PttHB2 genes are expressed in vascular tissue.

To confirm that the *PttHB* cDNAs are expressed in the cambial region, a northern blot analysis was performed on total RNA isolated from several different parts of the hybrid aspen plant. This showed that the full length *PttHB1* probe hybridised to RNA of xylem origin, less to RNA from leaf, much less to RNA from phloem and bark, and not at all to RNA from root (Figure 4). By contrast, the *PttHB2* probe gave a strong signal on both xylem and phloem RNA, a weaker signal on bark and leaf RNA, and a relatively very weak signal on root RNA.

The PttHB1 gene is expressed in the xylem maturation zone

To more precisely define the expression pattern of the *PttHB1* and *PttHB2* genes, an additional expression analysis was performed at higher resolution. To this end, tissues were isolated from different zones of the cambial region by cryo-microtome sectioning. PCR amplified total cDNAs, synthesised from mRNA isolated from these tissues, were hybridised to the short *PttHB1* probe. The exact anatomical location of the different tissue sections used for RNA isolation is shown in Figure 5A. As can be seen from Figure 5B, the expression of the *PttHB1* mRNA, as reflected by the amplified cDNAs, was clearly confined to the xylem.

In fact, the *PttHB1* gene was expressed in a single developmental zone, namely the maturing xylem (C4) (Figure 5B). Despite the PCR amplification step, only weak bands could be detected in the other cambial region zones. The *PttHB2* gene displayed a very different expression pattern from the *PttHB1* gene, being expressed mostly in the cambial (C2), enlarging xylem (C3) and enlarging phloem (C1) zones, and to a lesser extent in the maturing xylem zone (Figure 5B).

Because cells in the xylem maturation zone are undergoing a unique type of cell wall differentiation, we are especially intrigued by the *PttHB1* expression. During this phase, xylem cells are being mechanically strengthened by cell wall thickenings. Microfilaments re-orient and aid in determining the amount and pattern of the thickening. Cortical microtubules aid in positioning and depositing cellulose and lignin in the cell wall matrix. Ultimately, individual cells differentiate into one of several possible xylem cell types, each cell type in turn consisting of cells with differences in ultrastructure. Finally the xylem cells loose their cytoplasm and plasma membrane and die. It is very likely that the *PttHB1* HD protein is involved in the triggering of the specific developmental switch initiating the secondary wall formation in the xylem maturation zone.

Previous history has shown a high correlation between homeobox gene expression and the initiation of novel developmental pathways or regulatory circuits. Genetically identified mutant phenotypes in different developmental processes have also frequently been shown to be HB gene mutants. Similarly, when a HB gene has been isolated based on sequence similarity, and its molecular function later identified, in all cases it has encoded a key regulatory function in a defined developmental or environmentally induced pathway.

Preliminary assays on fibres isolated from transgenic plants by maceration in H_2O_2 and acetic acid at 100°C for four h in a Kajaani FiberLab apparatus showed that plants with sub normal levels *PttHB1* mRNA, as a result of an introduced antisense expression of the PttHB1 cDNA by gene technology, have slightly thinner cell walls (recorded as relative Kajaani CWT units) as compared to wt. This indicates that the isolated *PttHB* genes can be used as tools to modify fibrous secondary wall characteristics *in situ* already during their formation.

Although the invention has been described with regard to its preferred embodiments, which constitute the best mode presently known to the investigators, it should be understood that various changes and modifications as would be obvious to one having the ordinary skill in this art may be made without departing from the scope of the invention which is set in forth in the claims appended hereto.

Claims

1. A sequence class of homeobox genes (PALE) for regulating the fibre properties of fibrous plants, **characterised** in that proteins, encoded by genes belonging to said class exhibit a penta amino acid loop extension.

2. Sequence class according to claim 1, **characterised** in that said penta amino acid loop contains an amino acid QKIK motif.

3. Sequence class according to claim 1 or 2, **characterised** in that a characteristic WTP motif is included in the N-terminal arm of the homeodomain in most proteins of said PALE class, and an ITXE motif in the second helix of the structure, breaking this somewhat.

4. Sequence class according to claim 1 or 2, **characterised** in that DNA sequences belonging to that class exhibit at least a 50% identity with at least one of the following sequences:

SEQUENCE ID NO. 1 and SEQUENCE ID NO. 3.

5. Isolated DNA sequence regulating the fibre properties of fibrous plants, **characterised** in that said sequence exhibits at least a 50% identity with at least one of the following sequences:

SEQUENCE ID NO. 1 and SEQUENCE ID NO. 3.

6. Isolated DNA sequence regulating the fibre properties of fibrous plants, **characterised** in that said sequence is capable of hybridising to at least one of following:

SEQUENCE ID NO. 1 and SEQUENCE ID NO. 3.

7. Homeodomain protein or proteins regulating the cell differentiation of fibrous plants, **characterised** in that said protein/proteins exhibit at least 40% identity with at least one of the following aa sequences:

SEQUENCE ID NO. 2 and SEQUENCE ID NO. 4.

8. Use of a homeobox gene belonging to the sequence class (PALE) according to any one of claims 1 - 4 for the regulation of the fibre properties of a fibrous plant.

9. Use of a homeobox gene belonging to the sequence class (PALE) according to any one of claims 1 – 4 for the regulation of the fibre properties of a woody plant belonging to the group comprising coniferous (softwood) and dicotyledonous (softwood) trees.

10. Use of a DNA sequence according to any one of claims 5- 6 for the regulation of the fibre properties of a fibrous plant.

11. Use of a DNA sequence according to any one of claims 5- 6 for the regulation of the fibre properties of a woody plant belonging to the group comprising coniferous (softwood) and dicotyledonous (softwood) trees.

12. Use of a homeodomain protein according to claim 7, for the regulation of the fibre properties of a fibrous plant.

13. Use of a homeodomain protein according to claim 7, for the regulation of the fibre properties of a woody plant belonging to the group comprising coniferous (softwood) and dicotyledonous (softwood) trees.

14. A method of producing transgenic fibrous plants that produce fibre having altered properties, comprising the steps:

(a) constructing a plant expression vector which comprises in sense orientation a sequence from SEQUENCE ID NO. 1 or SEQUENCE ID NO. 3 and which will express that sense-oriented sequence when introduced into plant cells; or

(b) constructing a plant expression vector which comprises in antisense orientation a sequence from SEQUENCE ID NO. 1 or SEQUENCE ID NO. 3 and which will express that antisense-oriented sequence when introduced into plant cells; or

(c) constructing a plant expression vector carrying a sequence from SEQUENCE ID NO. 1 or SEQUENCE ID NO. 3 and which in other ways will directly change the expression of said sequence when introduced into plant cells;

(d) introducing the plant expression vector into a fibrous plant so that the sense-oriented sequence in the plant expression vector is expressed in the cambial region of the resulting transgenic plants to produce fibres having altered properties compared to corresponding fibres of untransformed plants; or

(e) introducing the plant expression vector into a fibrous plant so that the antisense-oriented sequence in the plant expression vector is expressed in the cambial region of the resulting transgenic plants to produce fibres having altered properties compared to corresponding fibres of untransformed plants

(f) introducing the plant expression vector into a fibrous plant so that the in other ways altered sequence in the plant expression vector is expressed in the cambial region of the resulting transgenic plants to produce fibres having altered properties compared to corresponding fibres of untransformed plants

(g) selecting transgenic plants of any one of step (d), (e) and (f) which exhibit altered fibre properties compared to those of untransformed plants; and

(h) propagating the transgenic plants of step (d), (e) and (f).

15. Transgenic fibrous plant, **characterised** in that it comprises at least one functionally inserted gene belonging to the class of homeobox genes according to any one of claims 1 - 4.

16. Transgenic fibrous plant according to claim 15, **characterised** in that said plant is selected from plants belonging to the group comprising coniferous (softwood) and dicotyledonous (softwood) trees.

15. Transgenic fibrous plant according to claim 15, **characterised** in that said plant is selected from plants belonging to the group comprising annual angiosperms.

```
                                                                136
1   CTAAACTGTGTCTAAAGTTACAACCATATGTCATAGAAAGTGAAACATTATTAGGTAAAGAAAAGGGAGAAAAAGGGGCTAGCATATATACCAC
                1▶MetSer•••

94  TGCCTGTCAAGAGGTAGGCAGAGTCAAGAGGAAGTGGAGGCTC ATG GAG GAG GGG AGG TTT CAG AAT GGT GGA GGG CTA
                                                 1▶Met Glu Glu Gly Arg Phe Gln Asn Gly Gly Gly Leu

173 GGT GTG AAA GTG ATG ACT GAT GAG CAA ATG GAG ATG TTG AGG AAG CAA ATC TCT GTC TAT GCC ACT ATT
 13▶Gly Val Lys Val Met Thr Asp Glu Gln Met Glu Met Leu Arg Lys Gln Ile Ser Val Tyr Ala Thr Ile

242 TGT GAG CAG CTT GTT GAG ATG CAC AAG GCT GTC TCT GCC CAA CAG GAC TTT GCT GGC ATG GGG CTT GGG
 36▶Cys Glu Gln Leu Val Glu Met His Lys Ala Val Ser Ala Gln Gln Asp Phe Ala Gly Met Gly Leu Gly

311 AAT CCG TAC TGT GAT CCA TTA CTG TCA TCT GCT GTC CAC AAG ATA GGG TCG AGG CAG CGG TGG ACA CCG
 59▶Asn Pro Tyr Cys Asp Pro Leu Leu Ser Ser Ala Val His Lys Ile Gly Ser Arg Gln Arg Trp Thr Pro

380 AAA CCA GCG CAA CTT CAA ATC CTT GAG CAG ATT TTC GAA CAA TGC AAT GCG ACT CCA GGC AGG CAG AAG
 82▶Lys Pro Ala Gln Leu Gln Ile Leu Glu Gln Ile Phe Glu Gln Cys Asn Ala Thr Pro Gly Arg Gln Lys

449 ATC AAA GAT ATA ACC AGG GAA CTT GCA CAA CAT GGC CAA ATT TCT GAA ACA AAT GTC TAC AAT TGG TTC
105▶Ile Lys Asp Ile Thr Arg Glu Leu Ala Gln His Gly Gln Ile Ser Glu Thr Asn Val Tyr Asn Trp Phe
                (528)
                SacI

518 CAA AAC AGG AGA GCT CGT TCA AAA AGA AAG CAA TCA GCC GTA GTG CCA AAC AAT GGA GAA TCT GAA ATG
128▶Gln Asn Arg Arg Ala Arg Ser Lys Arg Lys Gln Ser Ala Val Val Pro Asn Asn Gly Glu Ser Glu Met

587 GAG ACA GAT ATT GAG TCT TTA AAA GAG AAG AAG ACC AGG GCA GAA GAC AGC CAA CCT GAC GAA AAC ACC
151▶Glu Thr Asp Ile Glu Ser Leu Lys Glu Lys Lys Thr Arg Ala Glu Asp Ser Gln Pro Asp Glu Asn Thr

                                                                787
656 ACA CCG ATG GCT GAC CAT ATG TAC TTC AAC AGT CCA GAC ATA GGA TTT GAC CAG CTG ATG GGG AAG ATT
174▶Thr Pro Met Ala Asp His Met Tyr Phe Asn Ser Pro Asp Ile Gly Phe Asp Gln Leu Met Gly Lys Ile

725 GAA TCC CCA GGT CCA GGG AGC TGT ATT CCG TAC TGG CAG ATG GAG CAA TAT GAC TTG TTT GGA TGA TGTCT
197▶Glu Ser Pro Gly Pro Gly Ser Cys Ile Pro Tyr Trp Gln Met Glu Gln Tyr Asp Leu Phe Gly •••

796 CTTCCAGATCGGTAGAGATGCTCCGGTAGAAGCTTTCTGAGTTCAAGGACTCTGATATTGACAGTAATATAATAGAGCATTTATGCAGCGAAA

889 GATAACATGGTACACAGAGAACTTACATTCACAAGGCTCATTAGGGCTTAGGACTATCGTTGTGGCATGCATTTTACTGATTGCCTAATTCCA

982 ACCATCTGGGATATTTAAGATGAGTGAAATATCATGGTATTAGTGGTAGTATAAATTTTTATTTCGGATTATGTCAAGCTCTGTATTTYGTCC

1075 ATTTGATTGGATCATTATGGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
```

Fig. 1

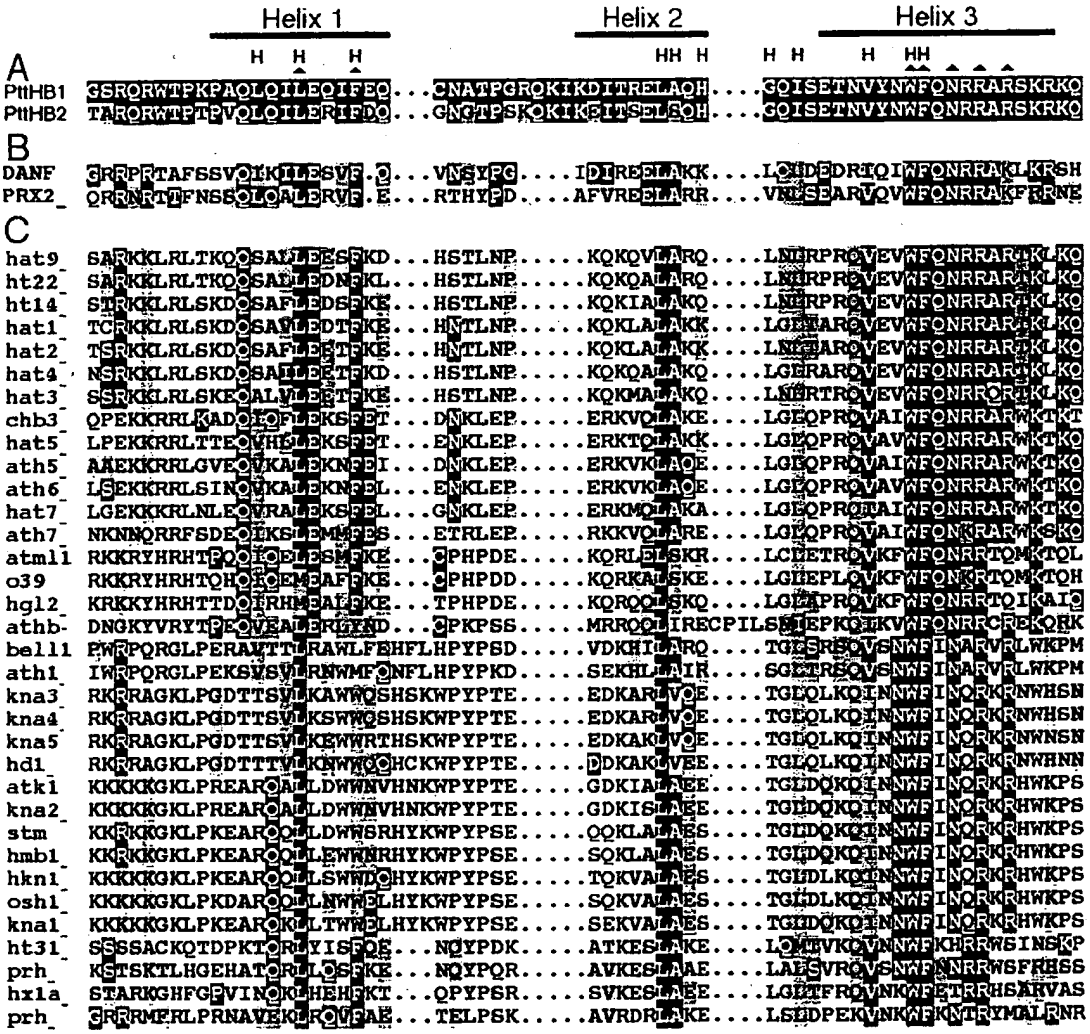

Helix 1
Helix 2
Helix 3
A
PttHB1
PttHB2
B
DANF
PRX2_
C
hat9_
ht22_
ht14_
hat1_
hat2_
hat4_
hat3_
chb3_
hat5_
ath5_
ath6_
hat7_
ath7_
atml1
o39
hgl2_
athb-
bell1
ath1_
kna3_
kna4_
kna5_
hd1_
atk1_
kna2_
stm_
hmb1_
hkn1_
osh1_
kna1_
ht31_
prh_
hx1a_
prh_

Fig. 2 A, B, C

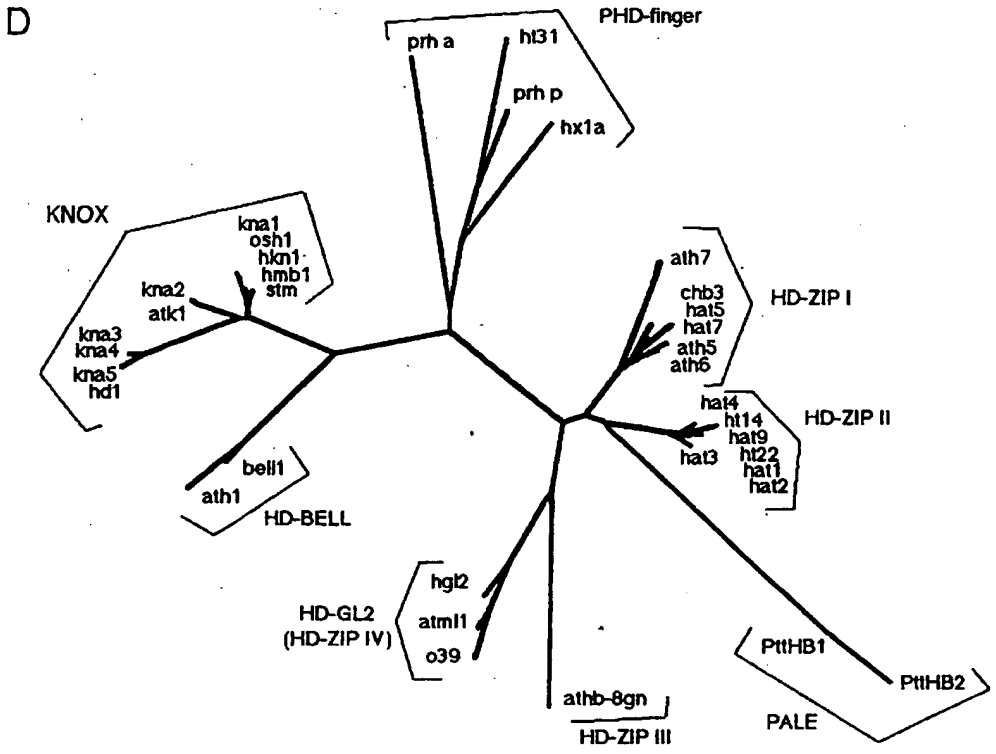

D
PHD-finger
prh a
ht31
prh p
hx1a
KNOX
kna1
osh1
hkn1
hmb1
stm
kna2
atk1
kna3
kna4
kna5
hd1
ath7
chb3
hat5
hat7
ath5
ath6
HD-ZIP I
hat4
ht14
hat9
hat3
ht22
hat1
hat2
HD-ZIP II
bell1
ath1
HD-BELL
hgl2
HD-GL2
(HD-ZIP IV)
atml1
o39
athb-8gn
HD-ZIP III
PttHB1
PttHB2
PALE

Fig. 2D

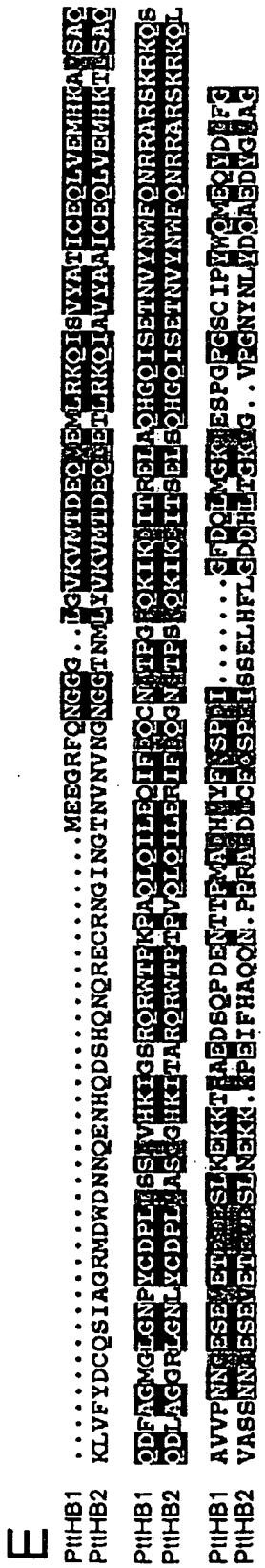


Fig. 2E

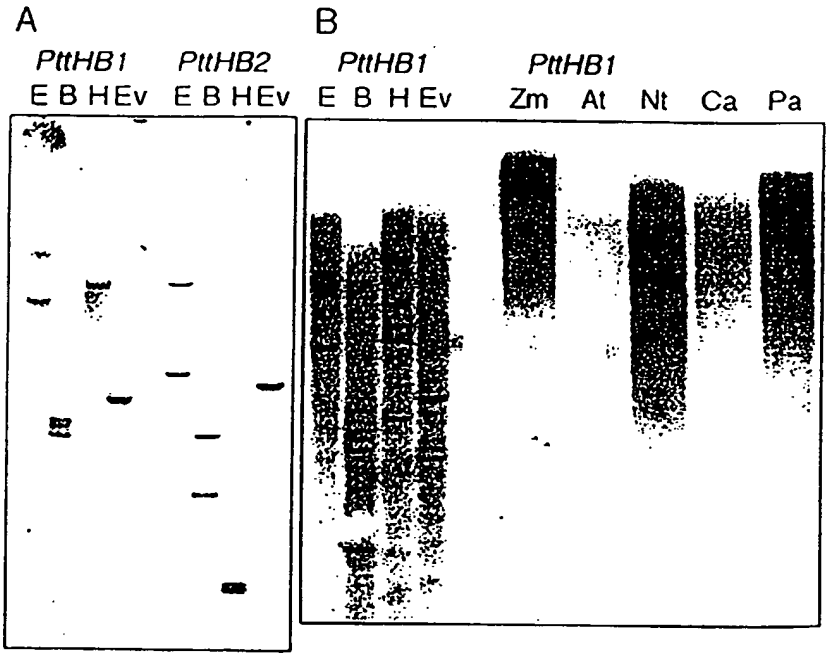

A
PttHB1
E B H Ev
PttHB2
E B H Ev
B
PttHB1
E B H Ev
PttHB1
Zm At Nt Ca Pa

Fig. 3

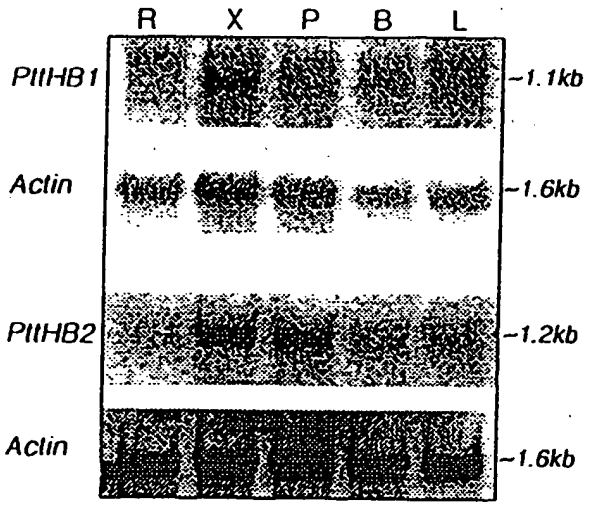


Fig. 4

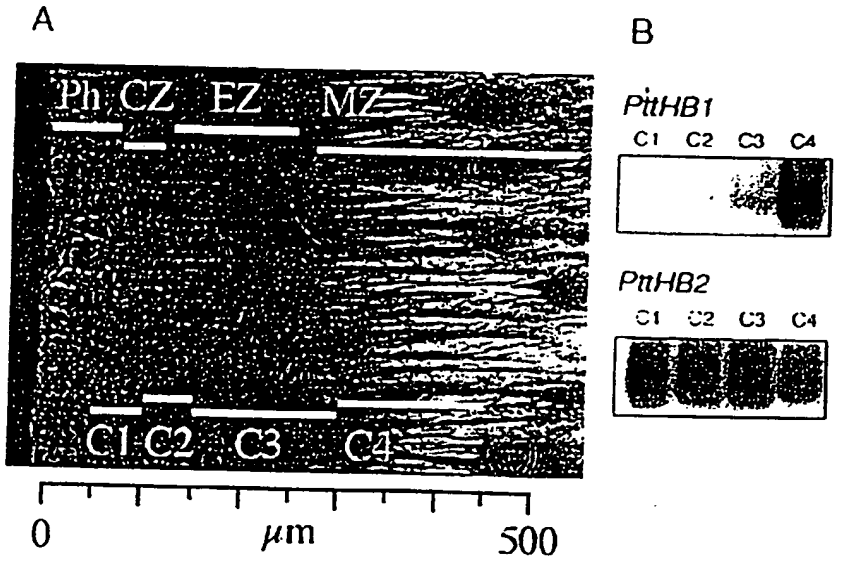

A
B
Ph CZ EZ MZ
C1 C2 C3 C4
0
µm
500
PttHB1
C1 C2 C3 C4
PttHB2
C1 C2 C3 C4

Fig. 5

SEQUENCE LISTING

<110> Hertzberg, Magnus
Olsson, Olof

<120> Novel sequence class of genes, corresponding amino acids and the use of the same

<130> PD53357PC00/MHG

<140>
<141>

<150> 9801129-9
<151> 1998-03-31

<160> 4

<170> PatentIn Ver. 2.1

<210> 1
<211> 1136
<212> DNA
<213> Populus tremula x Populus tremuloides

<300>
<301> Hertzberg, Magnus
Olsson, Olof
<302> Molecular characterisation of a novel plant homeobox gene expressed in the maturing xylem zone of Populus tremula x tremuloides
<303> Plant Journal
<304> 16
<306> 285-295
<307> 1998

<400> 1
ctaaactgtg tctaaagtta caaccatatg tcatagaaag tgaaacatta ttaggtaaag 60
aaaagggaga aaaggggcta gcatatatac cagtgcctgt caagaggtag gcagagtcaa 120
gaggaagtgg aggctcatgg aggaggggag gtttcagaat ggtggagggc taggtgtgaa 180
agtgatgact gatgagcaaa tggagatgtt gaggaagcaa atctctgtct atgccactat 240
ttgtgagcag cttgttgaga tgcacaaggc tgtctctgcc caacaggact ttgctggcat 300
ggggcttggg aatccgtact gtgatccatt actgtcatct gctgtccaca agatagggtc 360
gaggcagcgg tggacaccga aaccagcgca acttcaaatc cttgagcaga ttttcgaaca 420
atgcaatgcg actccaggca ggcagaagat caaagatata accagggaac ttgcacaaca 480
tggccaaatt tctgaaacaa atgtctacaa ttggttccaa aacaggagag ctcgttcaaa 540
aagaaagcaa tcagccgtag tgccaaacaa tggagaatct gaaatggaga cagatattga 600
gtctttaaaa gagaagaaga ccagggcaga agacagccaa cctgacgaaa acaccacacc 660

gatggctgac catatgtact tcaacagtcc agacatagga tttgaccagc tgatggggaa 720
gattgaatcc ccaggtccag ggagctgtat tccgtactgg cagatggagc aatatgactt 780
gtttggatga tgtctcttcc agatcggtag agatgctccg gtagaagctt tctgagttca 840
aggactctga tattgacagt aatataatag agcatttatg cagcgaaaga taacatggta 900
cacagagaac ttacattcac aaggctcatt agggcttagg actatcgttg tggcatgcat 960
tttactgatt gcctaattcc aaccatctgg gatatttaag atgagtgaaa tatcatggta 1020
ttagtggtag tataaatttt tatttcggat tatgtcaagc tctgtattty gtccatttga 1080
ttggatcatt atggaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaa 1136

<210> 2
<211> 217
<212> PRT
<213> Populus tremula x Populus tremuloides

<300>
<301> Hertzberg, Magnus
Olsson, Olof
<302> Molecular characterisation of a novel plant homeobox gene expressed in the maturing xylem zone of Populus tremula x tremuloides
<303> Plant Journal
<304> 16
<306> 285-295
<307> 1998

<400> 2

Met Glu Glu Gly Arg Phe Gln Asn Gly Gly Gly Leu Gly Val Lys Val
1 5 10 15

Met Thr Asp Glu Gln Met Glu Met Leu Arg Lys Gln Ile Ser Val Tyr
20 25 30

Ala Thr Ile Cys Glu Gln Leu Val Glu Met His Lys Ala Val Ser Ala
35 40 45

Gln Gln Asp Phe Ala Gly Met Gly Leu Gly Asn Pro Tyr Cys Asp Pro
50 55 60

Leu Leu Ser Ser Ala Val His Lys Ile Gly Ser Arg Gln Arg Trp Thr
65 70 75 80

Pro Lys Pro Ala Gln Leu Gln Ile Leu Glu Gln Ile Phe Glu Gln Cys
85 90 95

Asn Ala Thr Pro Gly Arg Gln Lys Ile Lys Asp Ile Thr Arg Glu Leu
100 105 110

Ala Gln His Gly Gln Ile Ser Glu Thr Asn Val Tyr Asn Trp Phe Gln

```
        115                 120                 125

Asn Arg Arg Ala Arg Ser Lys Arg Lys Gln Ser Ala Val Val Pro Asn
    130                 135                 140

Asn Gly Glu Ser Glu Met Glu Thr Asp Ile Glu Ser Leu Lys Glu Lys
145                 150                 155                 160

Lys Thr Arg Ala Glu Asp Ser Gln Pro Asp Glu Asn Thr Thr Pro Met
                165                 170                 175

Ala Asp His Met Tyr Phe Asn Ser Pro Asp Ile Gly Phe Asp Gln Leu
            180                 185                 190

Met Gly Lys Ile Glu Ser Pro Gly Pro Gly Ser Cys Ile Pro Tyr Trp
        195                 200                 205

Gln Met Glu Gln Tyr Asp Leu Phe Gly
    210                 215
```

<210> 3
<211> 1190
<212> DNA
<213> Populus tremula x Populus tremuloides

<300>
<301> Hertzberg, Magnus
Olsson, Olof
<302> Molecular characterisation of a novel plant homeobox gene expressed in the maturing xylem zone of Populus tremula x tremuloides
<303> Plant Journal
<304> 16
<306> 285-295
<307> 1998

<400> 3

```
aaagctagtg ttttatgact gccaaagtat agcaggcagg atggactggg acaacaatca 60
agaaaatcac caggacagtc atcaaaatca aagagagtgt agaaatggga ttaacgggac 120
taatgtgaat gtgaatggta atggtggaac taatatgttg tatgtgaaag ttatgacaga 180
tgaacaactg gaaactttaa ggaagcaaat tgcagtttat gctgctattt gtgaacagct 240
tgttgagatg cataaaaccc tctctgccca gcaagatctt gcaggaggca ggttgggaaa 300
tctatactgt gacccattga tggcatctgg gggtcacaaa attactgcca gacagcggtg 360
gactccaaca cctgtgcagc ttcaaattct tgagcggatc tttgatcaag gaaatggaac 420
tccaagcaag caaaagatta aggagataac atctgaactc agtcagcatg gacaaatttc 480
tgaaacaaat gtctataatt ggttccaaaa taggcgtgct cgatcaaaaa gaaagcagtt 540
agttgcatca tccaacaatg ctgaatcaga agtggagaca gaagttgatt cattaaatga 600
```

aaagaagaaa ccagaaatat ttcacgccca acagaaccct cctagggctg aagatctgtg 660
cttccaaagc cctgagataa gctctgaact tcatttcttg ggggatgacc acctgacagg 720
aaaggtggga gtgccaggga actataatct ttacgatcaa gcagaggact atggcatggc 780
gggctgactc tcattttgtc tccagatact aatttgtttg accaagttag atcttgacta 840
gaatccatgg ctttccaagt gtgaggttgt tggcgtgata gcaatgagaa cgtgaccttc 900
gctctaatta tggggttatt cttcacccat aaaagatgcc ttatatggtg cctgtcagtt 960
ggaagatgga tgaacctgta ctgtctcgca atttctattt ggtttgtgtt caatagattt 1020
accttcagga tcttttgtat gtgttcatag agtatacaaa tttgcggaga aaaaccgatg 1080
gtatgttaag ggtgtggagc ttttgctgga catgaactcg tgaagctgtc ttttttctga 1140
attcatgaag taaaattgtc tttaactttt gataaaaaaa aaaaaaaaaa 1190

<210> 4
<211> 261
<212> PRT
<213> Populus tremula x Populus tremuloides

<300>
<301> Hertzberg, Magnus
Olsson, Olof
<302> Molecular characterisation of a novel plant homeobox gene expressed in the maturing xylem zone of Populus tremula x tremuloides
<303> Plant Journal
<304> 16
<306> 285-295
<307> 1998

<400> 4
Lys Leu Val Phe Tyr Asp Cys Gln Ser Ile Ala Gly Arg Met Asp Trp
1 5 10 15

Asp Asn Asn Gln Glu Asn His Gln Asp Ser His Gln Asn Gln Arg Glu
20 25 30

Cys Arg Asn Gly Ile Asn Gly Thr Asn Val Asn Val Asn Gly Asn Gly
35 40 45

Gly Thr Asn Met Leu Tyr Val Lys Val Met Thr Asp Glu Gln Leu Glu
50 55 60

Thr Leu Arg Lys Gln Ile Ala Val Tyr Ala Ala Ile Cys Glu Gln Leu
65 70 75 80

Val Glu Met His Lys Thr Leu Ser Ala Gln Gln Asp Leu Ala Gly Gly
85 90 95

Arg Leu Gly Asn Leu Tyr Cys Asp Pro Leu Met Ala Ser Gly Gly His
100 105 110

```
Lys Ile Thr Ala Arg Gln Arg Trp Thr Pro Thr Pro Val Gln Leu Gln
       115                 120                 125

Ile Leu Glu Arg Ile Phe Asp Gln Gly Asn Gly Thr Pro Ser Lys Gln
    130                 135                 140

Lys Ile Lys Glu Ile Thr Ser Glu Leu Ser Gln His Gly Gln Ile Ser
145                 150                 155                 160

Glu Thr Asn Val Tyr Asn Trp Phe Gln Asn Arg Arg Ala Arg Ser Lys
                165                 170                 175

Arg Lys Gln Leu Val Ala Ser Ser Asn Asn Ala Glu Ser Glu Val Glu
            180                 185                 190

Thr Glu Val Asp Ser Leu Asn Glu Lys Lys Lys Pro Glu Ile Phe His
        195                 200                 205

Ala Gln Gln Asn Pro Pro Arg Ala Glu Asp Leu Cys Phe Gln Ser Pro
    210                 215                 220

Glu Ile Ser Ser Glu Leu His Phe Leu Gly Asp Asp His Leu Thr Gly
225                 230                 235                 240

Lys Val Gly Val Pro Gly Asn Tyr Asn Leu Tyr Asp Gln Ala Glu Asp
                245                 250                 255

Tyr Gly Met Ala Gly
            260
```

INTERNATIONAL SEARCH REPORT

International application No.
PCT/SE 99/00543

A. CLASSIFICATION OF SUBJECT MATTER

IPC6: C12N 15/29, A01H 5/00, A01H 7/00
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC6: C12N, A01H

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

SE,DK,FI,NO classes as above

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

BIOSIS, MEDLINE, STRAND, WPI, SCISEARCH

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P,X	The Plant Journal, Volume 16, No 3, 1998, Magnus Hertzberg et al, "Molecular characterisation of a novel plant homeobox gene expressed in the maturing xylem zone of Populus tremula tremuloides" page 285 - page 295 --	1-17
P,X	EMBL, Databas GenBank/DDBJ, accession no. AL031135, Koetter P. et al: "Arabidopsis thaliana DNA chromosome 4, BAC clone F8D20 (ESSAII project)"; & 1998-07-29, see nt. 19672-21025 --	1-7

[X] Further documents are listed in the continuation of Box C. [ ] See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" erlier document but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance: the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance: the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
29 June 1999	19 -07- 1999
Name and mailing address of the ISA/ Swedish Pat nt Offic Box 5055, S-102 42 STOCKHOLM Facsimile No. +46 8 666 02 86	Authorized officer Hampus Rystedt/Els Telephone No. +46 8 782 25 00

INTERNATIONAL SEARCH REPORT

International application No.
PCT/SE 99/00543

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	IAWA Journal, Volume 17, No 3, 1996, Rodney Arthur Savidge, "Xylogenesis, Genetic and Environmental Regulation -A Review-", page 269 - page 310, abstract, page 283 paragraph 3, page 286 paragraph 3 ---	1-17
T	Biochimica et Biophysica Acta, Volume 1442, 1998, Raquel L. Chan et al, "Homeoboxes in plant development" page 1 - page 19 --- --------	1-17